

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 11-282491

(43)Date of publication of application : 15.10.1999

(51)Int.CI.

G10L 3/00
 G10L 3/00
 G01C 21/00
 G08G 1/0969
 G09B 29/00
 G09B 29/10

(21)Application number : 10-084071

(71)Applicant : DENSO CORP

(22)Date of filing : 30.03.1998

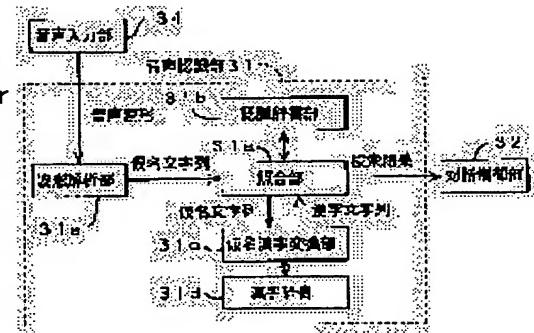
(72)Inventor : OKADE AKINORI

(54) SPEECH RECOGNITION DEVICE AND NAVIGATION SYSTEM

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a speech recognition device which can recognize a word written in Chinese characters as the one which a user intended, even if the user inputs the word in a pronunciation not registered in a dictionary means as kana (Japanese syllabary) character string data.

SOLUTION: A collation part 31a acquires a kana character string from a waveform analysis part 31e, and retrieves the data of the kana character string while collating it with the dictionary data stored in a recognition dictionary part 31b. And, if the retrieval is successful, the collation part outputs the recognition result to an interactive control part 32 and also notifies the success of the recognition thereto. On the other hand, the retrieval is unsuccessful, the collation part sends the kana character string data to a kana-kanji (Chinese character) conversion part 31c to let the kana-kanji conversion part 31c operate kana-kanji conversion using a kanji dictionary 31d, and acquires the conversion result. And, based on the converted kanji character-string data, the collation retrieval is operated by using the dictionary data stored in the recognition dictionary part 31b.



LEGAL STATUS

[Date of request for examination] 20.04.2004

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Best Available Copy

Copyright (C); 1998,2003 Japan Patent Office

(19) 日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平11-282491

(43) 公開日 平成11年(1999)10月15日

(51) Int.Cl. ⁶	識別記号	F I	
G 10 L 3/00	5 6 1	G 10 L 3/00	5 6 1 G
	5 5 1		5 6 1 D
			5 5 1 J
G 01 C 21/00		G 01 C 21/00	H
G 08 G 1/0969		G 08 G 1/0969	

審査請求 未請求 請求項の数 7 OL (全 13 頁) 最終頁に統ぐ

(21) 出願番号 特願平10-84071

(22) 出願日 平成10年(1998)3月30日

(71) 出願人 000004260

株式会社デンソー

愛知県刈谷市昭和町1丁目1番地

(72) 発明者 岡出 明紀

愛知県刈谷市昭和町1丁目1番地 株式会
社デンソー内

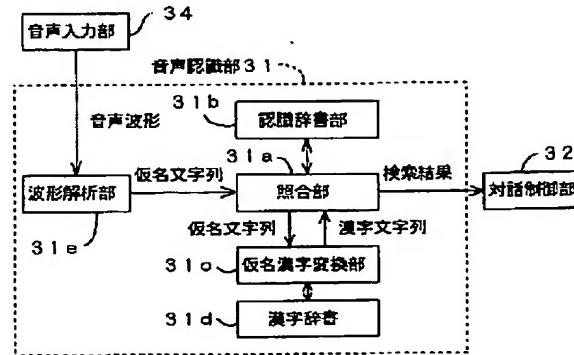
(74) 代理人 弁理士 足立 勉

(54) 【発明の名称】 音声認識装置及びナビゲーションシステム

(57) 【要約】

【課題】漢字表記される単語を、辞書手段に仮名文字列データとして登録されていない読み方で利用者が音声入力しても、利用者の意図する単語として認識可能な音声認識装置を提供する。

【解決手段】照合部31aは波形解析部31eから仮名文字列を取得し、その仮名文字列のデータに対して認識辞書部31b内に記憶されている辞書データとの照合を行いながら検索していく。そして、検索が成功であれば対話制御部32に対して、認識結果を出力すると共に認識成功を通知する。一方、検索が成功していないければ仮名文字列データを仮名漢字変換部31cに送り、その仮名漢字変換部31cにおいて漢字辞書31dを用いた仮名漢字変換させ、その変換結果を取得する。そして、その変換した漢字文字列データに基づき、認識辞書部31b内に記憶されている辞書データを用いて照合検索を行なう。



【特許請求の範囲】

【請求項1】 音声を入力するための音声入力手段と、該音声入力手段を介して入力された音声波形を解析して仮名文字列データに変換する波形解析手段と、該波形解析手段にて得た仮名文字列データを、予め辞書手段に記憶されている複数の比較対象データと比較して一致度合の高いものを認識結果とする認識手段と、を備える音声認識装置であって、前記辞書手段には、仮名文字列に対応する比較対象データと、当該仮名文字列を漢字に変換したものに対応する比較対象データとが記憶されており、さらに、仮名文字列と漢字文字列との変換を行う仮名漢字変換手段を備え、前記認識手段は、前記波形解析手段にて得た仮名文字列データを前記仮名漢字変換手段にて漢字文字列データに変換し、当該漢字文字列データを、前記辞書手段に記憶されている漢字文字列に対応する比較対象データと比較して一致度合の高いものを認識結果とすることが可能に構成されていること、を特徴とする音声認識装置。

【請求項2】 請求項1に記載の音声認識装置において、

前記認識手段は、

まず、前記波形解析手段にて得た仮名文字列データを前記辞書手段に記憶されている複数の比較対象データと比較して一致度合の高いものがあればそれを認識結果とし、

前記仮名文字列データでは一致度合いの高いものがない場合は、次に、前記仮名文字列データを漢字文字列データに変換し、前記辞書手段に記憶されている漢字文字列に対応する比較対象データと比較して一致度合の高いものがあればそれを認識結果とするよう構成されていること、

を特徴とする音声認識装置。

【請求項3】 請求項2に記載の音声認識装置において、

前記認識手段は、

まず、前記波形解析手段にて得た仮名文字列データを漢字文字列データに変換し、前記辞書手段に記憶されている漢字文字列に対応する比較対象データと比較して一致度合の高いものがあればそれを認識結果とし、

前記漢字文字列データでは一致度合いが高いものがない場合は、前記仮名文字列データを前記辞書手段に記憶されている複数の比較対象データと比較して一致度合の高いものがあればそれを認識結果とするよう構成されていること、

を特徴とする音声認識装置。

【請求項4】 請求項1～3のいずれかに記載の音声認識装置において、

さらに、

前記認識手段による認識結果を報知する報知手段と、該報知手段によって認識結果が報知された後に所定の確定指示がなされた場合には、当該認識結果を確定したものとして所定の確定後処理を実行する確定後処理手段と、を備えることを特徴とする音声認識装置。

【請求項5】 請求項4に記載の音声認識装置において、

前記報知手段は、音声を出力することにより報知する手段であり、前記認識結果の報知は、前記認識結果の内容を音声として前記音声出力手段から出力することによって行われることを特徴とする音声認識装置。

【請求項6】 請求項1～5のいずれかに記載の音声認識装置と、ナビゲーション装置とを備え、

前記音声認識装置の前記音声入力手段は、少なくとも前記ナビゲーション装置がナビゲート処理をする上で指定される必要のある所定のナビゲート処理関連データの指示を利用者が音声にて入力するために用いられるものであり、前記確定後処理手段は、前記認識手段による認識結果を前記ナビゲーション装置に出力するよう構成されていることを特徴とするナビゲーションシステム。

【請求項7】 請求項6に記載のナビゲーションシステムにおいて、

前記所定のナビゲート処理関連データは地名データであること、

を特徴とするナビゲーションシステム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、例えばナビゲーションシステムにおける目的地の設定などを音声によって入力できるようにする場合などに有効な音声認識装置及びその音声認識装置を備えたナビゲーションシステムに関する。

【0002】

【従来の技術】 従来より、入力された音声を予め記憶されている複数の比較対象データの候補と比較し、一致度合の高いものを認識結果とする音声認識装置が既に実用化されており、例えばナビゲーションシステムにおいて設定すべき目的地を利用者が地名を音声で入力するためなどに用いられている。特に車載ナビゲーションシステムを運転手自身が利用する場合、音声入力であればボタン操作や画面注視が伴わないとため、車両の走行中に行つても安全性が高いため有効である。

【0003】 このような機能を満たすため従来の音声認識装置においては、漢字で表現される言葉も認識可能とするため、漢字の「読み」を仮名で表したデータを辞書手段（データベース）に登録していた。この場合には、音声入力手段を介して入力された利用者の音声波形を波形解析手段にて解析して仮名文字列データに変換し、そ

の仮名文字列データを、辞書手段に記憶されている複数

の比較対象データと比較して一致度合の高いものを認識結果としていた。

【0004】

【発明が解決しようとする課題】ところで、上述した音声認識装置においては、辞書手段に登録されているとおりの読み方で利用者が入力しないと、装置側にてその内容を適切に認識できることとなる。例えばナビゲーションシステムにおいては目的地などを音声入力する場合には、読み方が幾通りも想定される地名があるが、この場合には辞書手段に登録されていない読み方では、所望の目的地が設定できないといった不都合が考えられる。そして、特に地名の場合にはその地方独特の読み方も多いため、この問題がより顕著となる。

【0005】具体的には、例えば「中原」という地名が九州地方にある。地元では「なかばる」と読むのが一般的であるが、それを知らない人は通常「なかはら」あるいは「なかばら」と読むことが想定される。また、「上本町」という地名が大阪にある。これは「うえほんまち」と読むのが一般的であるが、それを知らない人にとっては「うえもとまち」、「かみほんまち」、「かみもとまち」、「うえほんちょう」、「かみほんちょう」、「かみもとちょう」、「じょうほんちょう」などの読み方が考えられるため、どれが正しい読み方かが判らない可能性もある。したがって、これらの場合に「なかばる」あるいは「うえほんまち」と音声入力しない限り目的地を設定できないのでは、利便性などの点で問題である。

【0006】この場合の解決策として、例えば考えられる読み方について全て辞書手段に登録しておくことが考えられる。例えば「うえほんまち」だけでなく、「うえもとまち」、「かみほんまち」、「かみもとまち」、「うえほんちょう」、「かみほんちょう」、「かみもとちょう」、「じょうほんちょう」の全てについて地名の「上本町」に対応するような辞書データを構築するのである。

【0007】しかしながら、このように辞書データを構築すると、辞書データとして必要な記憶容量が増加すると共に、データベース作成のコストの増加にもつながり、好ましいことではない。なお、このような問題は、上述した地名には限らず、同じように複数の読み方が想定される単語を音声入力するような場面では同様に生じる。例えば、人名などについても同様である。したがって、このような問題を内在する単語が音声入力された場合に認識することを目的とする音声認識装置であれば同様の問題が生じることとなる。

【0008】本発明は、このような問題を解決し、漢字表記される単語を、辞書手段に仮名文字列データとして登録されていない読み方で利用者が音声入力しても、利用者の意図する単語として認識可能な音声認識装置を提供することを目的とする。

【0009】

【課題を解決するための手段及び発明の効果】上記目的を達成するためになされた請求項1に記載の音声認識装置によれば、利用者が音声入力手段を介して音声を入力すると、波形解析手段が、その入力された音声波形を解析して仮名文字列データに変換する。そして、認識手段は、波形解析手段にて得た仮名文字列データを予め辞書手段に記憶されている複数の比較対象データと比較して一致度合の高いものを認識結果とすることができる。また、辞書手段には、仮名文字列に対応する比較対象データに加え、その仮名文字列を漢字に変換したものに対応する比較対象データも記憶されている。さらに、仮名文字列と漢字文字列との変換を行う仮名漢字変換手段を備えている。そして認識手段は、仮名漢字変換手段にて仮名文字列データから変換された漢字文字列データを、辞書手段に記憶されている漢字文字列に対応する比較対象データと比較し、一致度合の高いものを認識結果とすることができるようにもされている。

【0010】このことによる効果を明確にするため、上20述した地名例「上本町」を用いて説明する。なお、実際には「大阪上本町」というように市のレベルから音声入力するのが現実的であるので、この例で説明する。辞書手段には「大阪上本町」の正しい読み方である「おおさかうえほんまち」という仮名文字列データと、それを漢字に変換した「大阪上本町」という漢字文字列データが記憶されている。例えば利用者が「大阪上本町」の正しい読み方である「おおさかうえほんまち」を知らずに「おおさかかみほんまち」と音声入力した場合には、当然、辞書手段に記憶されている「おおさかうえほんまち」という仮名文字列データとは一致しない。しかしながら、仮名漢字変換手段によって「おおさかかみほんまち」を漢字に変換すると「大阪上本町」という漢字文字列データを得ることができる。この「大阪上本町」という漢字文字列に対応する比較対象データは辞書手段に記憶されているので、この「大阪上本町」を認識結果とすることができる。なお、仮名漢字変換の際に例えば「大阪紙本町」というように「大阪上本町」以外の変換がされることもあるが、変換可能な漢字文字列に順次変換していくべき「大阪上本町」という漢字文字列は得られる。

【0011】このように、本発明の音声認識装置を用いれば、漢字表記される単語を、辞書手段に仮名文字列データとして登録されていない読み方で利用者が音声入力しても、利用者の意図する単語として認識することができる。そして、辞書手段には漢字文字列とその正しい読み方である仮名文字列とだけを記憶しておけばよいので、辞書手段の記憶容量を必要に増大させることがない。つまり、上述したように、「大阪上本町」に対してどのような読み方がされても利用者の意図する単語として認識するための他の解決策としては、考えられる読み方について全て、つまり「おおさかうえほんまち」だけ

でなく、「おおさかうえもとまち」、「おおさかかみほんまち」、「おおさかかみもとまち」、「おおさかうえほんちょう」、「おおさかかみほんちょう」、「おおさかかみもとちょう」、「おおさかじょうほんちょう」の全てについて地名の「大阪上本町」に対応するような辞書データを構築することが考えられる。しかし、このような辞書データの構築では、記憶容量が増加すると共に、データベース作成のコストの増加にもつながり、好みのことではない。そして、地名を正しく読める利用者にとっては、正しい読み方以外の読み方での仮名文字列データは不要なデータとなるため、一律にこのようなデータを辞書手段に持つことはやはり好みのことではない。その点からも、本発明の音声認識装置の有利さが伺える。

【0012】ところで、上述した効果は、認識手段が、仮名文字列データに基づく認識処理だけでなく、その仮名文字列データを変換した漢字文字列データに基づく認識処理も実行できる点から導かれるものであった。この場合、仮名文字列データに基づく認識処理を優先して実行してもよいし、逆に漢字文字列データに基づく認識処理を優先して実行してもよい。つまり、仮名文字列データに基づく認識処理を優先して実行する場合には、まず、仮名文字列データを辞書手段に記憶されている複数の比較対象データと比較して一致度合の高いものがあればそれを認識結果とし、仮名文字列データでは一致度合いの高いものがない場合は、次に、その仮名文字列データを漢字文字列データに変換し、辞書手段に記憶されている漢字文字列に対応する比較対象データと比較して一致度合の高いものがあればそれを認識結果とする。一方、漢字文字列データに基づく認識処理を優先して実行する場合には、まず、波形解析手段にて得た仮名文字列データを漢字文字列データに変換し、辞書手段に記憶されている漢字文字列に対応する比較対象データと比較して一致度合の高いものがあればそれを認識結果とし、漢字文字列データでは一致度合いが高いものがない場合は、仮名文字列データを辞書手段に記憶されている複数の比較対象データと比較して一致度合の高いものがあればそれを認識結果とする。

【0013】これら仮名あるいは漢字のいずれの文字列データに基づく認識処理を優先して実行するかについては、その認識対象などの要因でそれぞれメリット・デメリットがあるため、それらを考慮することが好みしい。そのメリット・デメリットについて説明する。

【0014】例えば、上述したように利用者が地名を音声入力する利用例で考えてみると、利用者が正しい地名を音声入力する可能性が高い状況においては、仮名文字列データに基づく認識処理を優先して実行する方が有効であると考える。つまり、この状況で漢字文字列データに基づく認識処理を優先すると、利用者が正しい地名を音声入力していても一律に仮名漢字変換処理が必要とな

ってくるのに対し、仮名文字列データに基づく認識処理を優先した場合には、利用者が正しい地名を音声入力している限り、仮名漢字変換処理及びその変換後の漢字文字列データに基づく認識処理を実行しなくてもよいため、最終的な認識結果を得るまでの速度が向上するからである。

【0015】一方、逆に、利用者が最初から正しい地名を音声入力する可能性が低い状況においては、漢字文字列データに基づく認識処理を優先して実行する方が有効であると考える。つまり、この状況で仮名文字列データに基づく認識処理を優先すると、その認識処理にて一致度合いが高いもののがなく、その後に仮名漢字変換処理を行い、その変換後の漢字文字列データに基づく認識処理を実行する可能性が高くなり、必ず認識処理を2回以上実行することとなるのに対し、漢字文字列データに基づく認識処理を優先した場合には、利用者が音声入力した地名に対してとにかく仮名漢字変換処理を行ってから認識処理を行うので、認識処理が1回で済む可能性が高くなり、相対的に、最終的な認識結果を得るまでの速度が向上するからである。

【0016】したがって、利用者の事情によりマッチした方の認識処理を優先して実行することが好みしい。なお、これらいずれから優先した認識処理を選択的に実行できるようにしておき、例えば同じ利用者であっても、地名の読み方をよく知らない状態では漢字文字列データに基づく認識処理を優先するように設定し、ある程度地名の正しい読み方を知った状態では、仮名文字列データに基づく認識処理を優先するように設定してもよい。

【0017】なお、認識手段による認識結果は報知手段によって報知し、認識結果が報知された後に所定の確定指示がなされた場合に、確定後処理手段が、その認識結果を確定したものとして所定の確定後処理を実行するように構成することも考えられる。そして、この認識結果の報知に関しては、所定の音声発生装置から認識結果の内容を音声にて出力することにより行うことが考えられる。カーナビゲーションシステムなどの車載機器用として用いる場合には、音声で出力されれば、ドライバーは視点を表示装置にずらしたりする必要がないので、安全運転のより一層の確保の点では有利であると言える。但し、音声出力に限定されるものではなく、画面上に文字または記号を表示できる表示装置に、認識結果の内容を、文字または記号による画像にて表示することにより行ったり、音声及び画像の両方にて報知するようにもよいし、それら以外の報知の手法を採用してもよい。車載機器として適用する場合に音声出力が有利であることを述べたが、もちろん車両が走行中でない状況もあるので、音声及び画像の両方で報知すれば、ドライバーは表示による確認と音声による確認との両方が可能となる。

【0018】なお、このように報知後に確定指示がなさ

れた場合にははじめて認識結果を確定するようにしたのは次の理由からである。つまり、利用者にとっては自分が音声入力した通りに装置に認識されているかどうか不安であるため、例えば「おおさかうえほんまち」と音声入力した場合に、「おおさかうえほんまちですね」と装置が応答してくれれば利用者自身で確認できるからである。そして、認識結果がそれでOKであれば、例えば「はい」と音声入力したり、あるいは操作スイッチにて確定を指示すればよい。

【0019】また、上述した音声認識装置をナビゲーションシステム用として用いる場合には、次に示すように構成することが考えられる。すなわち、上述した音声認識装置と、ナビゲーション装置とを備え、音声認識装置の音声入力手段は、少なくともナビゲーション装置がナビゲート処理をする上で指定される必要のある所定のナビゲート処理関連データの指示を利用者が音声にて入力するために用いられるものであり、確定後処理手段は、認識手段による認識結果をナビゲーション装置に出力するよう構成されているのである。この場合の「所定のナビゲート処理関連データ」としては、目的地などの地名が代表的なものとして挙げられる。

【0020】また、上述のナビゲーションシステムは、車載機器として用いられる場合だけではなく、例えば携帯型ナビゲーション装置や屋内用空調装置などでもよい。但し、これまで説明したように車載機器用として用いる場合には利用者がドライバーであることが考えられ、その場合には運転自体が最重要であり、それ以外の車載機器については、なるべく運転に支障がないことが好ましい。したがって、車載機器としてのナビゲーションシステムを前提とした音声認識装置の場合には、より一層の利点がある。

【0021】なお、上述した音声認識装置やそれを用いたナビゲーションシステムにおいては、利用者によって音声入力される対象として地名を例に挙げて説明したが、地名には限らず、同じように複数の読み方が想定される単語、例えば人名などについて音声入力し、それを認識する装置として実現した場合にも、同様の効果がある。

【0022】

【発明の実施の形態】以下、本発明が適用された実施例について図面を用いて説明する。なお、本発明の実施の形態は、下記の実施例に何ら限定されることなく、本発明の技術的範囲に属する限り、種々の形態を探り得ることは言うまでもない。

【0023】図1は本実施例の音声認識装置30を適用したカーナビゲーションシステム2の概略構成を示すブロック図である。本カーナビゲーションシステム2は、位置検出器4、地図データ入力器6、操作スイッチ群8、これらに接続された制御回路10、制御回路10に接続された外部メモリ12、表示装置14及びリモコン

センサ15及び音声認識装置30を備えている。なお制御回路10は通常のコンピュータとして構成されており、内部には、周知のCPU、ROM、RAM、I/O及びこれらの構成を接続するバスラインが備えられている。

【0024】前記位置検出器4は、いずれも周知のジャイロスコープ18、距離センサ20、及び衛星からの電波に基づいて車両の位置を検出するGPS(Global Positioning System)のためのGPS受信機22を有している。これらのセンサ等18、20、22は各々が性質の異なる誤差を持っているため、複数のセンサにより、各々補間しながら使用するように構成されている。なお、精度によっては上述した内の一部で構成してもよく、更に、ステアリングの回転センサ、各転動輪の車輪センサ等を用いてもよい。

【0025】地図データ入力器6は、位置検出の精度向上のためのいわゆるマップマッチング用データ、地図データ及び目印データを含む各種データを入力するための装置である。媒体としては、そのデータ量からCD-ROMを用いるのが一般的であるが、メモリカード等の他の媒体を用いても良い。

【0026】表示装置14はカラー表示装置であり、表示装置14の画面には、位置検出器4から入力された車両現在位置マークと、地図データ入力器6より入力された地図データと、更に地図上に表示する誘導経路や後述する設定地点の目印等の付加データとを重ねて表示することができる。

【0027】また、本カーナビゲーションシステム2は、リモートコントロール端末(以下、リモコンと称する。)15aを介してリモコンセンサ15から、あるいは操作スイッチ群8により目的地の位置を入力すると、現在位置からその目的地までの最適な経路を自動的に選択して誘導経路を形成し表示する、いわゆる経路案内機能も備えている。このような自動的に最適な経路を設定する手法は、ダイクストラ法等の手法が知られている。操作スイッチ群8は、例えば、表示装置14と一緒にになったタッチスイッチもしくはメカニカルなスイッチ等が用いられ、各種入力に使用される。

【0028】そして、音声認識装置30は、上記操作スイッチ群8あるいはリモコン15aが手動操作により目的地などを指示するために用いられるのに対して、利用者が音声で入力することによっても同様に目的地などを指示することができるようにするための装置である。

【0029】この音声認識装置30は、「波形解析部」、「認識手段」としての音声認識部31と、対話制御部32と、音声合成部33と、音声入力部34と、「音声入力手段」としてのマイク35と、PTT(Push-To-Talk)スイッチ36と、スピーカ37とを備えている。

【0030】音声認識部31は、音声入力部34から入

力された音声データを、対話制御部32からの指示により入力音声の認識処理を行い、その認識結果を対話制御部32に返す。すなわち、音声入力部34から取得した音声データに対し、記憶している辞書データを用いて照合を行ない、複数の比較対象パターン候補と比較して一致度の高い上位比較対象パターンを対話制御部32へ出力する。入力音声中の単語系列の認識は、音声入力部34から入力された音声データを順次音響分析して音響的特徴量（例えばケプストラム）を抽出し、この音響分析によって得られた音響的特徴量時系列データを得る。そして、周知のDFTマッチング法によって、この時系列データをいくつかの区間に分け、各区間が辞書データとして格納されたなどの単語に対応しているかを求める。

【0031】対話制御部32は、その認識結果及び自身が管理する内部状態から、音声合成部33への応答音声の発声指示や、システム自体の処理を実行する制御回路10に対して例えばナビゲート処理のために必要な目的地を通知して設定処理を実行させるよう指示する処理を実行する。このような処理が確定後処理であり、結果として、この音声認識装置30を利用すれば、上記操作スイッチ群8あるいはリモコン15aを手動しなくとも、音声入力によりナビゲーション装置に対する目的地の指示などが可能となるのである。

【0032】また音声入力部34は、マイク35にて取り込んだ周囲の音声をデジタルデータに変換して音声認識部31に出力するものである。本実施例においては、利用者がPTTスイッチ36を押しながらマイク35を介して音声を入力するという使用方法である。具体的には、音声入力部34はPTTスイッチ36が押されたかどうかを判断しており、PTTスイッチ36が押されている場合にはマイク35を介しての音声入力処理を実行するが、押されていない場合にはその音声入力処理を実行しないようにしている。したがって、PTTスイッチ36が押されている間にマイク35を介して入力された音声データのみが音声認識部31へ出力されることとなる。

【0033】ここで、音声認識部31についてさらに説明する。図2は、この音声認識部31の構成をさらに詳しく示したものである。図2に示すように、音声認識部31は、照合部31a、認識辞書部31b、仮名漢字変換部31c、漢字辞書部31d及び波形解析部31eとを備えている。波形解析部31eは、音声入力部34から取得した音声波形を解析して仮名文字列データに変換する。そして、照合部31aは、その仮名文字列のデータに基づき、認識辞書部31b内に記憶されている仮名文字列による辞書データを用いて照合を行なうことができる。さらに照合部31aは、波形解析部31eから得た仮名文字列データを仮名漢字変換部31cに送り、その仮名漢字変換部31cに漢字辞書31dを用いた仮名漢字変換をさせて得た漢字文字列データに基づき、認識

辞書部31b内に記憶されている漢字文字列による辞書データを用いて照合を行なうこともできるようにされている。

【0034】そして、照合部31aにて複数の比較対象データと比較されて一致度が高いとされた上位比較対象パターンは、対話制御部32へ出力される。ところで、本実施例の認識辞書部31bに記憶されている辞書データは、仮名文字列に対応する比較対象データだけではなく、その仮名文字列を漢字に変換したものに対応する比較対象データも記憶されている。具体的には、地名データとして、地名が漢字で表される場合（漢字を一部に含む場合も同様）には、その漢字文字列のデータと、その地名の正しい読み方を仮名文字列で表したデータとがセットで記憶されている。つまり、大阪の上本町であれば、「大阪上本町」という漢字文字列データと、その正しい読み方である「おおさかうえほんまち」とが記憶されている。

【0035】なお、仮名漢字変換部31cやその仮名漢字変換部31cが参照する漢字辞書31dは、一般的なものでよく、例えば日本語ワープロ用の仮名漢字変換機能及び漢字辞書の内容を援用したものでも構わない。上述したように、音声認識部31においては、照合部31aが、波形解析部31eから取得した仮名文字列データそのもの、あるいはその仮名文字列データを仮名漢字変換部31cにて仮名漢字変換させて得た漢字文字列データに基づき、認識辞書部31b内に記憶されている辞書データを用いて照合を行なうことができる。そして、一致度の高い上位比較対象データを対話制御部32へ出力する。

【0036】対話制御部32では、図示しない記憶部にこの上位比較対象データを記憶しておき、所定の確定指示がなされた場合に制御回路10へデータを送って所定の処理をするように指示する「確定後処理」を実行したり、あるいは音声合成部33へ音声データを送って発音させるように指示する処理を実行する。

【0037】次に、本実施例のカーナビゲーションシステム2の動作について説明する。なお、音声認識装置30に関係する部分が特徴があるので、カーナビゲーションシステムとしての一般的な動作を簡単に説明した後、音声認識装置30に関係する部分の動作について詳しく説明することとする。

【0038】カーナビゲーションシステム2の電源オン後に、表示装置14上に表示されるメニューから、ドライバーがリモコン15a（操作スイッチ群8でも同様に操作できる。以後の説明においても同じ）により、案内経路を表示装置14に表示させるために経路情報表示処理を選択した場合、あるいは、音声認識装置30を介して希望するメニューをマイク35を介して音声入力することで、対話制御部32から制御回路10へ、リモコン15aを介して選択されるのと同様の指示がなされた場

合、次のような処理を実施する。

【0039】すなわち、ドライバーが表示装置14上の地図に基づいて、音声あるいはリモコンなどの操作によって目的地を入力すると、GPS受信機22から得られる衛星のデータに基づき車両の現在地が求められ、目的地と現在地との間に、ダイクストラ法によりコスト計算して、現在地から目的地までの最も短距離の経路を誘導経路として求める処理が行われる。そして、表示装置14上の道路地図に重ねて誘導経路を表示して、ドライバーに適切なルートを案内する。このような誘導経路を求める計算処理や案内処理は一般的に良く知られた処理であるので説明は省略する。

【0040】次に、音声認識装置30における動作について、上述の経路案内のための目的地を音声入力する場合を例にとって説明する。図3は、その場合の音声認識部31及び対話制御部32における処理を示すフローチャートである。なお、本フローチャートは、音声認識部31が図2に示す構成を備えていることを前提とした処理を示している。

【0041】まず最初のステップS10においては音声入力があるかどうかを判断する。上述したように、PTTスイッチ36が押されている間にマイク35を介して入力された音声データのみが音声入力部34から音声認識部31へ出力されるので、この音声入力部34からの音声入力があるかどうかを判断する。

【0042】音声入力があれば、S20へ移行して音声認識処理を実行する。この音声認識処理については後で詳述することとして、続くS30では、S20での音声認識処理によって認識が成功したかどうかを判断する。認識が成功していれば(S30: YES)、S40に移行し、S20での音声認識処理による認識結果が、「はい」という音声入力であるかどうかを判断する。そして、「はい」という音声入力でなければ(S40: NO)、続くS50にて認識結果が所定カテゴリに属するものであるかどうかを判断する。ここでは経路案内のための目的地を設定する処理を前提としているので、この所定カテゴリとは、地名に関するカテゴリである。

【0043】この所定カテゴリであれば(S50: YES)、S60へ移行して、その認識結果を音声にて応答する処理を実行する。これは、音声合成部33及びスピーカを介して認識結果を音声として出力する処理である。一方、所定カテゴリでなければ(S50: NO)、S80へ移行してその他の処理を実行する。S60あるいはS80の処理の後はS10へ戻って、処理を繰り返す。

【0044】また、S30にて否定判断、すなわち認識成功でなければ、S70へ移行し、認識失敗を音声にて応答する。さらに、S40にて肯定判断、すなわち認識結果が「はい」という音声入力であった場合には、S10へ移行して認識結果を確定する。そして続くS11

0にて、所定の確定後処理を実行する。この場合の確定後処理とは、認識結果としての「経路案内のための目的地」に関するデータを、制御回路10へ(図1参照)へ出力する処理などとなる。このような確定後処理が終了した後は、S10へ戻る。

【0045】次に、上記S20での音声認識処理の詳細について図4を参照して説明する。なお、図4のフローチャートは、音声認識部31の照合部31aを主体とした処理手順を示している。まず、照合部31aは波形解析部31eから仮名文字列を取得する(S201)。そして、その仮名文字列のデータに対して認識辞書部31b内に記憶されている辞書データとの照合を行いながら検索していく(S202)。そして、検索が成功したかどうかを判断し(S203)、検索が成功であれば(S203: YES)、S207へ移行する。S207では、対話制御部32に対して、認識結果を出力すると共に認識成功を通知する。

【0046】一方、検索が成功していないければ(S203: NO)、S204へ移行し、S201で取得した仮名文字列データを仮名漢字変換部31cに送り、その仮名漢字変換部31cにおいて漢字辞書31dを用いた仮名漢字変換をさせ、その変換結果を取得する(S204)。そして、その変換結果に基づき、仮名漢字変換の候補となる漢字があるかどうかを判断し(S205)、候補漢字があれば(S205: YES)、その漢字文字列データに基づき、認識辞書部31b内に記憶されている辞書データを用いて照合検索を行なう(S202)。また、候補漢字がなければ(S205: NO)、認識が失敗したことを対話制御部32へ通知してから(S206)、本処理ルーチンを終了する。

【0047】以上が、経路案内のための目的地を音声入力する場合を例にとっての動作説明であるが、本発明の音声認識に係る特徴及び効果をより明確に理解するために、目的地として「大阪上本町」を指定するという具体例で説明を続ける。まず、利用者がマイク35を介して正しい読み方である「おおさかうえほんまち」と音声入力した場合を考える。この場合、図4のS202において認識辞書部31b内を検索すると、対応する辞書データが存在するので検索成功となり(S203: YES)、S207で認識結果を出し、認識成功を通知する。そのため、図3のS30では肯定判断となるが、S40では否定判断されるため、S50へ移行する。この場合は地名に関するカテゴリであるので(S50: YES)、S60へ以降して、認識結果を音声応答する。

【0048】ここで例えば「おおさかうえほんまちですか」と音声応答され、利用者が「はい」と音声入力すると、今度は図3のS40にて肯定判断となるためS10へ移行し、認識結果を確定してからS110にて所定の確定後処理を実行する。この確定後処理とは、上述したように認識結果としての「おおさかうえほんまち」を

経路案内のための目的地として制御回路10へ(図1参照)へ出力する処理である。

【0049】一方、利用者がマイク35を介して間違った読み方である「おおさかかみほんちょう」と音声入力した場合を考える。この場合、図4のS202において認識辞書部31b内を検索すると、対応する辞書データが存在しないため検索失敗となり(S203: NO)、S204にて漢字辞書31dを用いた仮名漢字変換がされる。ここで候補漢字として「上本府」という漢字文字列に変換された場合には、候補漢字としては存在するため(S205: YES)、S202へ戻って認識辞書部31b内を検索することとなるが、この場合には検索が成功しないので、S204にて仮名漢字再変換を行う。

【0050】この仮名漢字再変換によって漢字文字列「大阪上本町」と変換された場合には、候補漢字として存在するため(S205: YES)、S202に戻って認識辞書部31b内を検索すると、今度は検索が成功するため(S203: YES)、S207で認識結果を出力し、認識成功を通知する。そのため、図3のS30では肯定判断となるが、S40では否定判断されるため、S50へ移行する。この場合は地名に関するカテゴリがあるので(S50: YES)、S60へ以降して、認識結果を音声応答する。

【0051】このS60においても、やはり「おおさかうえほんまちですか」と音声応答する。利用者は「おおさかかみほんちょう」と音声入力しているが、認識辞書部31b内には「大阪上本町」と「おおさかうえほんまち」とがセットで記憶されているため、「大阪上本町」で検索した結果成功したのであれば、その読み方である「おおさかうえほんまち」で音声応答する。利用者は自分の読み方が間違っていたことを知り、「はい」と音声入力すると、今度は図3のS40にて肯定判断となるためS100へ移行し、認識結果を確定してからS110にて所定の確定後処理を実行する。上述したように認識結果としての「おおさかうえほんまち」を経路案内のための目的地として制御回路10へ(図1参照)へ出力する処理である。

【0052】なお、この例では、最初の仮名漢字変換においては「上本府」と変換され、2回目に「大阪上本町」と変換されるように説明したが、最初の仮名漢字変換で「大阪上本町」と変換される場合もあるし、3回目以降になって初めて「大阪上本町」と変換されることもある。但し、仮名漢字変換部31cにいわゆる「学習機能」を持たせれば、前回「大阪上本町」と変換されていれば、今回の仮名漢字変換の際には最初の変換で「大阪上本町」と変換されることとなる。

【0053】また、仮名だけで表記されている地名の場合には、利用者が読み間違えるという自体を想定しなくてもよいため、図4のS202において認識辞書b内を検索すれば、必ず対応する辞書データが存在するため検

索成功となる(S203: YES)。一方、仮名と漢字が混じった地名の場合には、仮名漢字変換の際に、仮名混じり漢字文字列と変換されることで対応できる。

【0054】このように、本実施例の音声認識装置30を用いれば、例えば経路設定のための目的地を音声入力する際、漢字表記される地名を、認識辞書部31bに仮名文字列データとして登録されていない読み方で利用者が音声入力しても、利用者の意図する単語として認識することができる。そして、認識辞書部31bには漢字文字列(例えば「大阪上本町」とその正しい読み方である仮名文字列(例えば「おおさかうえほんまち」とだけを記憶しておけばよいので、認識辞書部31bの記憶容量を不必要に増大させることができない。つまり、上述したように、地名「大阪上本町」に対してどのような読み方がされても利用者の意図する単語として認識するための他の解決策としては、考えられる全ての読み方、つまり「おおさかうえほんまち」だけでなく、「おおさかうえもとまち」、「おおさかかみほんまち」、「おおさかみもとまち」、「おおさかうえほんちょう」、「おおさかかみもとちょう」、「おおさかじょうほんちょう」などの全てについて地名「大阪上本町」に対応するような辞書データを構築することが考えられる。しかし、このような辞書データの構築では、認識辞書部31bの記憶容量が増加すると共に、データベース作成のコストの増加にもつながり、好みないことではない。そして、地名を正しく読める利用者にとっては、正しい読み方以外の読み方での仮名文字列データは不要なデータとなるため、一律にこのようなデータを認識辞書部31bに持つことはやはり好みのことではない。その点からも、本実施例の音声認識装置30は有効である。

【0055】なお、経路設定のための目的地ではなく、単にその地区周辺の地図を表示させたい場合であっても、同様である。ところで、本実施例の音声認識装置30においては、仮名文字列データに基づく認識処理を優先して実行していた。つまり、図4に示すように、まず仮名文字列に基づいて認識辞書部31b内の辞書データとの照合検索を行い(S201, S202)、検索が成功しなければ(S203: NO)、次に、その仮名文字列を漢字文字列に変換し、変換された漢字文字列に基づいて認識辞書部31b内の辞書データとの照合検索を行うようしている(S204, S202)。

【0056】これに対して、逆に漢字文字列データに基づく認識処理を優先して実行することも考えられる。この場合の音声認識処理の具体的な処理手順例について図5を参照して説明する。波形解析部31eから仮名文字列を取得し(S501)、まず、その取得した仮名文字列データを仮名漢字変換部31cに送って、仮名漢字変換部31cにおいて漢字辞書31dを用いた仮名漢字変換をさせ、その変換結果を取得する(S502)。そし

て、その変換結果に基づき、仮名漢字変換の候補となる漢字があるかどうかを判断し(S 503)、候補漢字があれば(S 503: YES)、その漢字文字列データに基づき、認識辞書部31b内に記憶されている辞書データを用いて照合検索を行なう(S 504)。検索が成功であれば(S 505: YES)、S 506へ移行し、対話制御部32に対して、認識結果を出力すると共に認識成功を通知する。また、S 504での検索が成功でない場合には(S 505: NO)、S 502へ戻り、再度、仮名漢字変換を行ってその変換結果を得て、同様に照合検索を行う。

【0057】一方、S 503にて否定判断、つまり、S 501で得た仮名文字列を変換可能な全ての漢字文字列に変換して照合検索しても検索が成功しなかった場合には、S 501で取得した仮名文字列そのものに基づき、認識辞書部31b内に記憶されている辞書データを用いて照合検索を行なう(S 508)。検索が成功であれば(S 509: YES)、S 506へ移行し、対話制御部32に対して、認識結果を出力すると共に認識成功を通知する。また、S 508での検索が成功でない場合には(S 509: NO)、認識が失敗したことを対話制御部32へ通知してから(S 510)、本処理ルーチンを終了する。

【0058】なお、S 508では仮名文字列そのものに基づいて照合検索を行なっているが、それは次の理由からである。つまり、地名であっても漢字表記ではなく仮名表記しかないものもある。その場合に、仮名漢字変換(S 502)の結果、候補漢字がない場合に(S 503: YES)、そのまま何ら対処しないのでは、仮名表記しかない地名を認識できないこととなってしまう。したがって、対応する漢字文字列がない仮名文字列だけの地名についてはS 503で否定判断されるため、その場合には、仮名文字列そのものに基づいて照合検索しているのである。

【0059】この別処理例のように、最初に仮名漢字変換を行って漢字文字列に基づく照合検索を優先して実行する場合には、例えば「おおさかかみほんまち」や「おおさかうえほんちょう」など、上本町の正しい読み方である「おおさかうえほんまち」以外の読み方で利用者が音声入力した場合であっても、仮名漢字変換にてどの場合には「大阪上本町」という地名の漢字文字列が得られるため、それに基づく検索が成功することとなる。

【0060】このように、図4に示す音声認識処理では、仮名文字列データに基づく認識処理を優先して実行した場合を説明し、図5に示す音声認識処理では、漢字文字列データに基づく認識処理を優先して実行した場合を説明した。これら仮名あるいは漢字のいずれの文字列に基づく認識処理を優先して実行するかについては、その認識対象などの要因でそれぞれメリット・デメリットがあるため、それらを考慮することが好ましい。

【0061】つまり、利用者が正しい地名を音声入力する可能性が高い状況においては、図4に示すように仮名文字列データに基づく認識処理を優先して実行する方が有効である。なぜなら、この状況で図5に示すような漢字文字列データに基づく認識処理を優先すると、利用者が正しい地名を音声入力していても一律に仮名漢字変換処理(図5のS 502参照)が必要となってくるのに対し、仮名文字列データに基づく認識処理を優先した場合には、利用者が正しい地名を音声入力している限り、仮名漢字変換処理及びその変換後の漢字文字列データに基づく認識処理(図4のS 204, S 205, S 202参照)を実行しなくてもよいため、最終的な認識結果を得るまでの速度が向上するからである。

【0062】一方、逆に、利用者が最初から正しい地名を音声入力する可能性が低い状況においては、図5に示すように漢字文字列データに基づく認識処理を優先して実行する方が有効であると考える。つまり、この状況で図4に示すような仮名文字列データに基づく認識処理を優先すると、その仮名文字列に基づく検索が失敗して

20 (S 203: NO)、その後に仮名漢字変換処理(S 204)を行い、その変換後の漢字文字列データに基づく認識処理(S 202)を実行する可能性が高くなり、必ず認識処理(S 202)を2回以上実行することとなる。

それに対し、図5に示すような漢字文字列データに基づく認識処理を優先した場合には、利用者が音声入力した地名に対してとにかく仮名漢字変換処理(S 502)を行ってから認識処理(S 504)を行うので、認識処理(S 504)が1回で済む可能性が高くなり、相対的に、最終的な認識結果を得るまでの速度が向上するからである。

【0063】したがって、利用者の事情によりマッチした方の認識処理を優先して実行することが好ましい。そのため、音声認識装置30を構成する際に、いずれかに固定して実行する構成としてもよいが、両方の処理を選択的に行実行できるように構成しておいてもよい。そして、例えば同じ利用者であっても、地名の読み方をよく知らない状態では漢字文字列データに基づく認識処理を優先するように設定し、ある程度地名の正しい読み方を知った状態では、仮名文字列データに基づく認識処理を優先するように設定できるようになることが考えられる。これは、例えばリモコンセンサ15から、あるいは操作スイッチ群8によりモード設定を変更できるようにしておけば容易に対応できる。

【0064】以上、一実施例を説明したが、その他、種々の態様での変更が可能である。例えば、上記実施例では「報知手段」としてスピーカ37を用い、音声出力により認識結果を報知するようにしたが、このように音声で出力されれば、認識結果の確認のためにドライバーが視点を移動する必要がないので、一層の安全運転に貢献できる。つまり、カーナビゲーションシステム2を車載

機器用として用いているので、このような音声出力には利点がある。もちろん、画面上に文字または記号を表示することにより認識結果を報知してもよいし、音声を出力することにより報知すると共に画面上に文字または記号を表示することにより報知するようにしててもよい。そして、画面上に認識結果を表示させる場合には、カーナビゲーションシステム2の地図情報を表示するための表示装置14(図1参照)に表示させるような構成を採用することもできる。

【0065】なお、上記図3のS50の処理においては、認識結果が所定カテゴリに属するものであるかどうかを判断するものとし、その所定カテゴリとは目的地の設定を前提にするため地名に関するカテゴリであると説明した。しかしながら、本発明の主旨はこのような地名等に限定されるものではなく、抽象的に言えば、認識結果を出力し、利用者の確認を得てから正式に確定する必要があるような情報に関するカテゴリということである。具体的に上述のカーナビゲーションシステム2で言うならば、ナビゲート処理をする上で指定される必要のある所定のナビゲート処理関連情報の指示ということとなる。この「所定のナビゲート処理関連情報」の代表的なものが目的地であるが、それ以外にもルート探索に関する条件選択など、ナビゲート処理をする上で指定の必要な指示が含まれる。

【0066】また、上述した実施例では、音声認識装置30をカーナビゲーションシステム2に適用した例として説明したが、適用先としては、上述したカーナビゲーションシステム2には限定されない。例えば携帯型ナビゲーション装置であってもよいし、あるいは街頭やバーキングエリアなどに設置されて道案内などをするための情報端末装置などにも同様に適用できる。

【0067】さらに、上記実施例では、ナビゲーションシステム2における経路設定などのための目的地を音声入力するような場合を想定し、従って認識対象が地名であったが、認識対象としては地名には限らない。例えば、同じように複数の読み方が想定される単語、例えば人名などについて音声入力し、それを認識する装置として実現した場合にも、同様の効果がある。つまり、音読

みと訓読みの組み合わせなどによって同じ漢字表記の氏あるいは名であっても、複数の読み方ができる。したがって、例えば正しい読み方が判らなくても、とりあえず訓読みで音声入力すれば、それに対応する漢字文字列は得られるからである。

【図面の簡単な説明】

【図1】 本発明の実施例としてのカーナビゲーションシステムの概略構成を示すブロック図である。

【図2】 音声認識装置における音声認識部の構成を示すブロック図である。

【図3】 音声認識装置における音声認識及び対話制御に係る処理を示すフローチャートである。

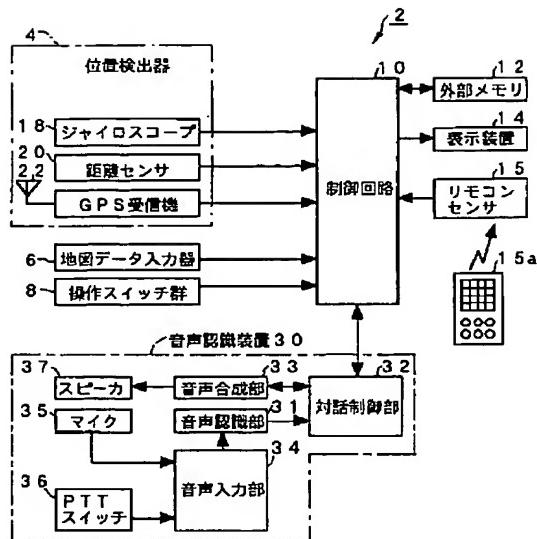
【図4】 音声認識装置における音声認識部にて実行される音声認識処理ルーチンを示すフローチャートである。

【図5】 音声認識装置における音声認識部にて実行される音声認識処理ルーチンの別処理例を示すフローチャートである。

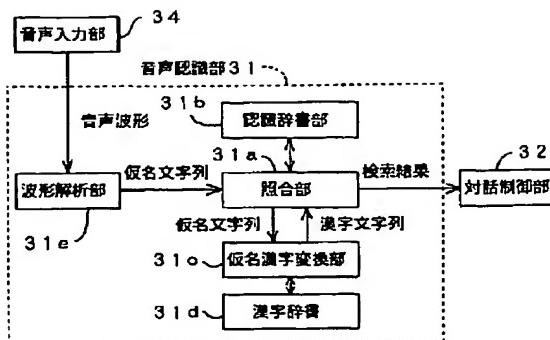
【符号の説明】

20	2…カーナビゲーションシステム 6…地図データ入力器 群	4…位置検出器 8…操作スイッチ ンサ 1 0…制御回路 1 4…表示装置 ンサ 1 5 a…リモコン スコープ 2 0…距離センサ 機	1 2…外部メモリ 1 5…リモコンセ 1 8…ジャイロ 2 2…G P S受信 機
30	3 0…音声認識装置 3 1 a…照合部 3 1 c…仮名漢字変換部 3 1 e…波形解析部 3 3…音声合成部 3 5…マイク ッチ 3 7…スピーカ	3 1…音声認識部 3 1 b…辞書部 3 1 d…漢字辞書部 3 2…対話制御 3 4…音声入力部 3 6…P T Tスイ	

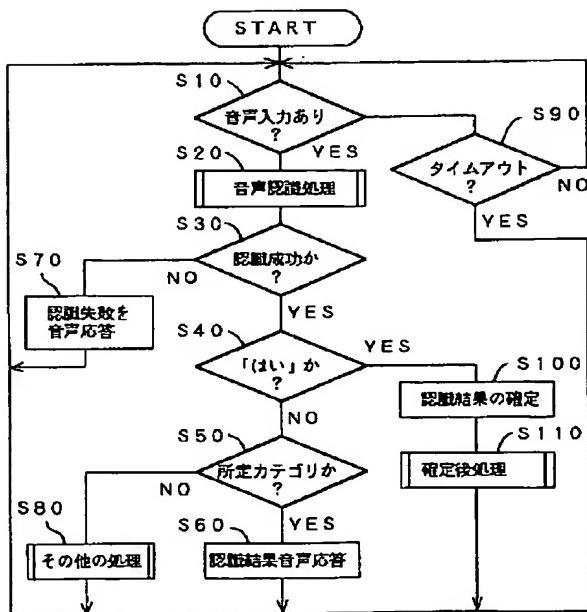
【図1】



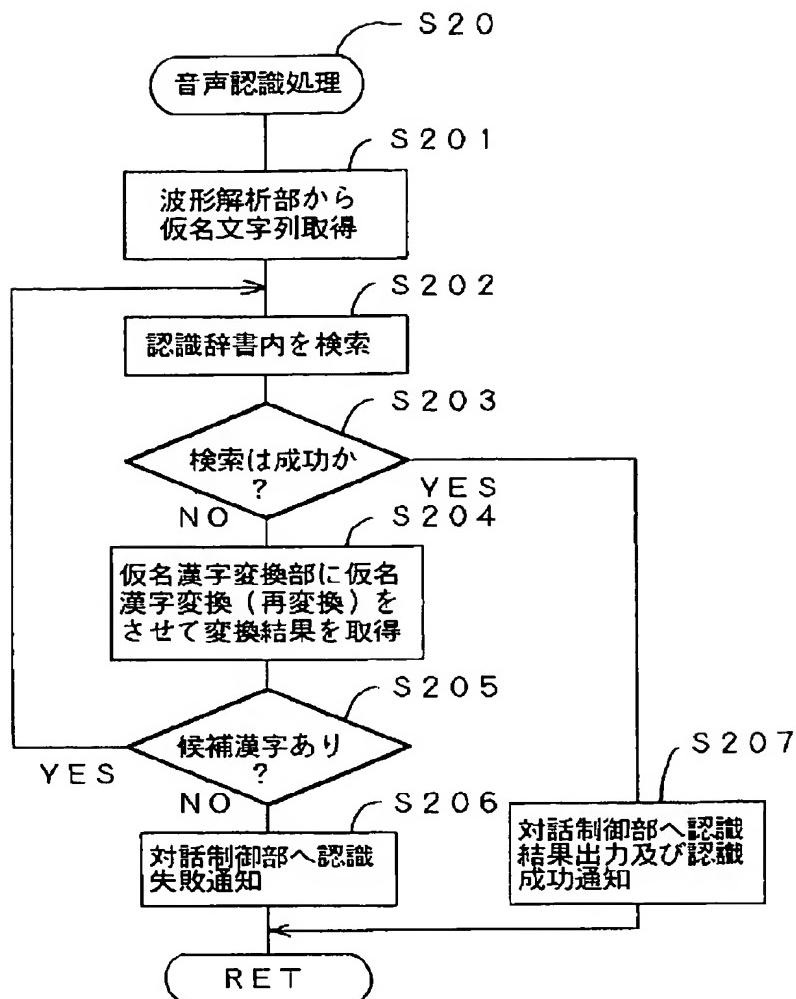
【図2】



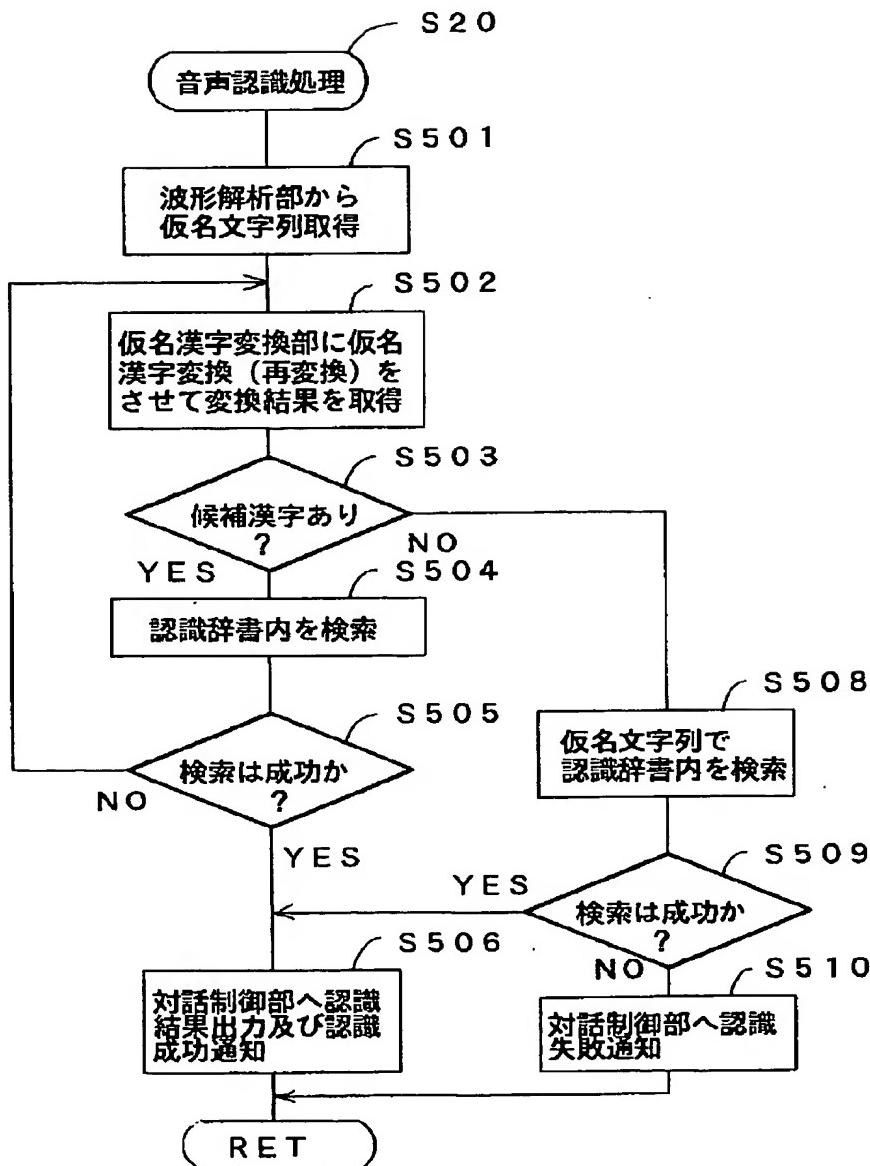
【図3】



【図4】



【図5】



フロントページの続き

(S1)Int.C1.^b
G 09 B 29/00
29/10

識別記号

F I
G 09 B 29/00
29/10

F
A

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER: _____**

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.